

19 BUNDESREPUBLIK
DEUTSCHLAND



DEUTSCHES
PATENTAMT

12 Patentschrift
10 DE 196 36 028 C 1

51 Int. Cl.⁶:
G 06 K 9/62
G 08 G 1/04
G 06 T 7/20

21 Aktenzeichen: 196 36 028.5-53
22 Anmeldetag: 5. 9. 96
43 Offenlegungstag: —
45 Veröffentlichungstag
der Patenterteilung: 20. 11. 97

DE 19636028 C 1

Innerhalb von 3 Monaten nach Veröffentlichung der Erteilung kann Einspruch erhoben werden

73 Patentinhaber:

Daimler-Benz Aktiengesellschaft, 70567 Stuttgart,
DE

72 Erfinder:

Franke, Uwe, Dr.-Ing., 73066 Udingen, DE

56 Für die Beurteilung der Patentfähigkeit
in Betracht gezogene Druckschriften:

DE 44 31 479 A1
DE 43 08 776 A1
EP 06 26 655 A2

SANEYOSHI Keiji »3-D image recognition system by
means of stereoscopy combined with ordinary
image processing«, IEEE Symposium on Intelli- gent
Vehicles '94, 24.10.1994 bis 26.10.1994, Paris, S.13-18;
L. MATTHIES et al. »Obstacle detection for
unmanned ground vehicles: a progress report«,
IEEE Symposium on Intelligent Vehicles '95,
25./26.09.1995, Detroit, S.66-71;

54 Verfahren zur Stereobild-Objektdetektion

57 Die Erfindung bezieht sich auf ein Verfahren zur Detektion
und gegebenenfalls Bewegungsverfolgung von Objekten
mittels Stereobildauswertung.

Erfindungsgemäß wird aus dem jeweils aufgenommenen
Stereobildpaar ein Strukturklassenbildpaar erzeugt, wobei
für jeden Bildpunkt die Helligkeitsdifferenzen ausgewählter
Umgebungsbildpunkte als Digitalwerte ermittelt und letztere
zu einer jeweiligen Digitalwertgruppe zusammengefaßt wer-
den, wobei jeweils identische Gruppen eine eigene Struktur-
klasse definieren. Die Strukturklassen mit längs der Epipol-
arlinie fehlender Helligkeitsänderung werden verworfen. Für
die Bildpunkte der übrigen Strukturklassen werden zugehö-
rige Disparitätswerte bestimmt und mit einem gegebenen
Häufigkeitsinkrement in einem Disparitätshistogramm akku-
muliert. Die zu einem jeweiligen Häufungspunktbereich des
Histogramms gehörige Bildpunktgruppe wird dann als ein zu
detektierendes Objekt interpretiert.

Verwendung z. B. zur Objektdetektion im Straßenverkehr als
Hilfsmittel zur autonomen Längs- oder Querverführung von
Fahrzeugen.

DE 19636028 C 1

Die Erfindung bezieht sich auf ein Verfahren zum Detektieren und gegebenenfalls auch zum Verfolgen von Objekten mittels Aufnahme und Auswertung von Stereobildern. Ein solches Verfahren ist beispielsweise nützlich als Hilfsmittel zum automatisierten Führen von Straßenfahrzeugen oder auch auf dem Gebiet der Robotik, wo es jeweils darauf ankommt, relevante Objekte zu detektieren und deren Position und Größe zu bestimmen. Im Fall der autonomen Fahrzeugführung, z. B. im städtischen Stop&go-Verkehr, sind neben stehenden Objekten, die sich im Fahrbereich des Fahrzeugs befinden, auch alle bewegten Objekte im näheren Umfeld zu detektieren. Hingegen ist für derartige Anwendungen keine vollständige, dichte Tiefenkarte der Szene erforderlich. So kann meist von einer relativ ebenen Grundfläche und deutlich erhabenen Objekten ausgegangen werden. Es brauchen zudem keine Detailinformationen, wie die genaue Form eines vorausfahrenden Fahrzeugs, abgeleitet werden. Derartige vereinfachende Randbedingungen gelten analog für eine Vielzahl von Problemstellungen auf den Gebieten der Robotik und der Überwachungstechnik.

Die Verfahren zur Stereobild-Objektdetektion lassen sich in flächenbasierte Methoden und merkmalsbasierte Methoden unterteilen. Flächenbasierte Methoden sind z. B. in den Konferenzbeiträgen K. Sane Yoshi, 3-D image recognition system by means of stereoscopy combined with ordinary image processing, Intelligent Vehicles '94, 24.10.1994 bis 26.10.1994, Paris, Seiten 13 bis 18 und L. Matthies et al, Obstacle detection for unmanned ground vehicles: a progress report, Intelligent Vehicles '95, 25./26.09.1995, Detroit, Seiten 66 bis 71 beschrieben. Sie erfordern einen höheren Rechenaufwand als merkmalsbasierte Methoden. Für einen Überblick über gängige Stereobild-Objektdetektionsverfahren sei auf das Buch O. Faugeras, Three-Dimensional Computer Vision, MIT-Press, 1993 verwiesen.

Aus der Offenlegungsschrift DE 44 31 479 A1 ist ein Verfahren zur Detektion von Objekten, insbesondere Fahrzeugen, bekannt; bei dem von einem jeweiligen Bereich zwei Bilder aus verschiedenen Blickwinkeln aufgenommen werden. Aus einem Vergleich der beiden Bilder, insbesondere deren Grauwerte, wird für mindestens einen Teilbereich jeweils dann ein Objekt erkannt, wenn der Unterschied zwischen den beiden Bildern für den betreffenden Teilbereich größer als eine vorgegebene Schwelle ist.

Bei einem in der Offenlegungsschrift DE 43 08 776 A1 offenbarten System zur Überwachung eines fahrzeugexternen Zustandes ist eine Stereobild-Objektdetektion vorgesehen, durch die ein jeweiliges Objekt innerhalb einer festen Region außerhalb eines Fahrzeugs abgebildet wird. Die aufgenommenen Bilder werden einer Bildverarbeitungsvorrichtung zugeführt, welche eine Distanzverteilung über das gesamte Bild hinweg berechnet. Um ein jeweiliges Objekt im linken und rechten Bild aufzufinden, wird das jeweilige Bild in kleine Regionen unterteilt, und es werden Farb- oder Helligkeitsmuster innerhalb dieser Regionen für die beiden Bilder verglichen, um die Regionen mit sich entsprechenden Objektdetails herauszufinden und daraus die Distanzverteilung über das gesamte Stereobild zu bestimmen.

In der Offenlegungsschrift EP 0 626 655 A2 ist eine Einrichtung zur Erkennung vorausfahrender Fahrzeuge und zur Bestimmung von deren Entfernung beschrieben, die von einem Stereobild-Objektdetektionsverfah-

ren Gebrauch macht. Für die Bildauswertung werden spezielle, auf das Erkennen von Fahrzeugkonturen zugeschnittene Techniken eingesetzt, welche Fahrzeugkonturmodelle verwenden. Für ein erkanntes Fahrzeug wird die zugehörige Disparität und daraus die Fahrzeugentfernung ermittelt.

Der Erfindung liegt als technisches Problem die Bereitstellung eines Verfahrens zur Stereobild-Objektdetektion zugrunde, das sich besonders für die oben genannten Anwendungsfälle mit vereinfachten Randbedingungen eignet und vergleichsweise effizient und robust ist.

Die Erfindung löst dieses Problem durch die Bereitstellung eines Verfahrens mit den Merkmalen des Anspruchs 1. Bei diesem Verfahren wird aus dem aufgenommenen Stereobildpaar zunächst ein Strukturklassenbildpaar erzeugt, indem für jeden Bildpunkt die Helligkeitsdifferenzen zu vorgegebenen Umgebungsbildpunkten in digitaler Form ermittelt und die Digitalwerte in einer vorbestimmten Reihenfolge zu einer Digitalwertgruppe zusammengefaßt werden. Jede der verschiedenen möglichen Gruppen definiert dabei eine eigene Strukturklasse. Anschließend wird eine vergleichsweise effektiv durchführbare, zweckmäßig einfach gestaltete Korrespondenzanalyse durchgeführt, bei der alle Strukturklassen unberücksichtigt bleiben, die in Richtung der Epipolarlinie, d. h. zur Linie sich entsprechenden Bildpunktpaare eines jeweiligen gemeinsamen Urbildpunktes, keinen Strukturgradienten zeigen.

Dies sind diejenigen Strukturklassen, deren Bildpunkte sich in ihrer Helligkeit nicht um ein vorgegbares Maß von der Helligkeit der in Richtung der Epipolarlinie liegenden Umgebungsbildpunkte unterscheiden. Dadurch wird bereits ein merklicher Bildverarbeitungsaufwand eingespart. In dieser Richtung verlaufende Strukturen eignen sich systembedingt bei der merkmalsbasierten Stereobildauswertung ohnehin nicht zur Entfernungsbestimmung. Zu allen übrigen Strukturklassen werden dann Disparitätswerte korrespondierender Bildpunkte gleicher Strukturklasse ermittelt, die in einem Disparitätshistogramm zu einem jeweiligen Häufigkeitswert akkumuliert werden. Optional kann für jedes korrespondierende Bildpunktpaar der zugehörige Disparitätswert mit einem gewichteten Häufigkeitsinkrement in das Histogramm eingehen. Anschließend wird das Histogramm auf Häufungsbereiche untersucht. Zu festgestellten, interessierenden Häufungsbereichen wird dann umgekehrt die zugehörige Bildpunktgruppe eines zugehörigen Strukturklassenbildes dargestellt und als in einer bestimmten Entfernung befindliches Objekt interpretiert. Es zeigt sich, daß dieses Objektdetektionsverfahren für viele Anwendungsfälle sehr effektiv arbeitet und robuste, d. h. wenig störanfällige Resultate liefert.

Sei einem nach Anspruch 2 weitergebildeten Verfahren wird eine ternäre Logik zur Digitalisierung der Helligkeitsdifferenzen verwendet, was eine für die hier betrachteten Anwendungen sehr vorteilhafte, weil einerseits ausreichend differenzierte und andererseits ausreichend schnell durchführbare Strukturklassifizierung ermöglicht.

Bei einem nach Anspruch 3 weitergebildeten Verfahren werden als Umgebungsbildpunkte die vier Bildpunkte gewählt, die an den jeweiligen Bezugspunkt zum einen parallel und zum anderen senkrecht zur Epipolarlinie jeweils beidseits direkt angrenzen oder von diesem um eine vorgebbare Abgriffweite beabstandet angeordnet sind. Eine Abgriffweite in der Größe eines

oder eventuell auch mehrerer Bildpunkte ermöglicht die Einbeziehung eines größeren Nachbarschaftsbereiches, was in Fällen zweckmäßig ist, in denen sich die Helligkeitsgradienten typischerweise über einen Bereich mehrerer Bildpunkte hinweg erstrecken.

Bei einem nach Anspruch 4 weitergebildeten Verfahren ist der erforderliche Berechnungsaufwand weiter dadurch verringert, daß kleine Disparitäten favorisiert werden, indem zu jedem Bildpunkt des einen Strukturklassenbildes nur die minimale Disparität, d. h. die Entfernung des nächstgelegenen Bildpunktes gleicher Strukturklasse im anderen Strukturklassenbild, ermittelt und für die Erstellung des Disparitätshistogramms berücksichtigt wird. Die zugrundeliegende Annahme der Favorisierung kleiner Disparitäten und damit großer Objektentfernungen ist insbesondere bei der Verwendung in Fahrzeugen für die Betrachtung von Straßenszenen mit ihren relativ großen Hindernissen angemessen. Mit dieser Vorgehensweise wird zudem das Auftreten von nahen Phantomobjekten bei der Korrespondenzanalyse unterdrückt, wie sie durch weit entfernte, periodische Strukturen, z. B. Häuserfronten, generiert werden können.

Bei einem nach Anspruch 5 weitergebildeten Verfahren wird jedem Bildpunkt neben der die Helligkeitsdifferenzen repräsentierenden Digitalwertgruppe ein die Größe der aufgetretenen Helligkeitsdifferenzen stärker differenzierender Kontrastwert zugeordnet, der als Gewichtungsfaktor zur Bildung eines gewichteten Häufigkeitsinkrements verwendet wird. Für jedes korrespondierende Bildpunktpaar, für das ein Disparitätswert ermittelt wird, wird dann selbiger mit dem kontrastabhängig gewichteten Häufigkeitsinkrement zum Disparitätshistogramm akkumuliert. Dadurch werden kontrastreiche Strukturen stärker bei der Bildauswertung berücksichtigt, was das Grundrauschen vermindert.

Bei einem nach Anspruch 6 weitergebildeten Verfahren wird die Objektdetektion durch eine Ballungsanalyse der als zu einem jeweiligen Objekt gehörig interpretierten Bildpunktgruppe erweitert, die einerseits ausreichend einfach gestaltet ist und andererseits eine verbesserte Lokalisierung und Bewegungsverfolgung des Objektes erlaubt.

Bei einem nach Anspruch 7 weitergebildeten Verfahren wird anschließend an die Detektion eines Objektes eine kontinuierliche Bewegungsverfolgung desselben durchgeführt, indem die erfindungsgemäße Objektdetektion beschränkt auf einen minimalen, das detektierte Objekt enthaltenden Kubus zyklisch wiederholt durchgeführt wird. Die Beschränkung auf den hierfür jeweils relevanten Bildbereich spart nicht nur Rechenaufwand ein, sondern reduziert vor allem auch den Einfluß störender Hintergrundobjekte. Außerdem braucht nur noch ein relativ kleiner Bereich von Disparitätswerten ausgewertet werden, wobei die zugehörige Tiefe des Kubus von der maximal zu erwartenden Entfernungsänderung bestimmt ist.

Ein nach Anspruch 8 weitergebildetes Verfahren eignet sich für Straßenfahrzeuge und beinhaltet ein plausibles Straßenmodell, das es gestattet, Strukturen der Straßenoberfläche mit hoher Zuverlässigkeit herauszufiltern, was die Detektion von Objekten im Bereich über der Straßenoberfläche erleichtert und/oder das zuverlässige Erfassen von Fahrbahnmarkierungen zur autonomen Fahrzeugquerführung ermöglicht.

Eine vorteilhafte Ausführungsform der Erfindung ist in den Zeichnungen illustriert und wird nachfolgend beschrieben. Hierbei zeigen:

Fig. 1 ein von einer Stereokameraanordnung eines Fahrzeugs aufgenommenes Stereobildpaar,

Fig. 2 eine schematische Darstellung der zur Auswertung der Stereobilder gemäß Fig. 1 verwendeten Strukturklassifizierung,

Fig. 3 eine Grauwertdarstellung des durch die Strukturklassifizierung gemäß Fig. 2 aus dem linken Stereobild von Fig. 1 erhaltenen Strukturklassenbildes,

Fig. 4 ein durch Korrespondenzanalyse der beiden, zu den Stereobildern von Fig. 1 gehörigen Strukturklassenbildern erhaltenes Disparitätshistogramm mit mehreren Häufungspunkten,

Fig. 5 eine Darstellung der mit einem ersten Häufungspunkt von Fig. 4 verknüpften Bildpunktgruppe und

Fig. 6 eine Darstellung der mit einem zweiten Häufungspunkt von Fig. 4 verknüpften Bildpunktgruppe.

Fig. 1 zeigt ein Stereobildpaar 1a, 1b, das eine typische Straßenverkehrsszene wiedergibt. Zur kontinuierlichen Aufnahme des Verkehrsgeschehens vor einem Fahrzeug anhand solcher Stereobildpaare 1a, 1b ist an dem Fahrzeug in herkömmlicher Weise eine, nach vorn gerichtete Stereokameraanordnung angebracht, deren optische Achsen vorzugsweise parallel zueinander ausgerichtet und horizontal gegeneinander versetzt sind. In diesem Fall bilden die zugehörigen Epipolarlinien, d. h. diejenigen Linien, entlang denen der zu einem gemeinsamen Urbildpunkt gehörige Bildpunkt im einen Stereobild gegenüber dem entsprechenden Bildpunkt im anderen Stereobild versetzt sein kann, gleichzeitig jeweilige Bildzeilen des Stereo-Bildpaares 1a, 1b, was die spätere Disparitätsauswertung besonders einfach macht. Die Stereobildüberwachung des Verkehrsgeschehens vor dem Fahrzeug läßt sich insbesondere zum automatisierten Führen des Fahrzeugs in Längs- und/oder Querrichtung bzw. zur Warnung vor auftretenden Hindernissen, wie anderen Fahrzeugen und Fußgängern, verwenden. Interessierende Objekte sind bei dem Stereobildpaar 1a, 1b von Fig. 1 folglich vor allem das auf derselben Fahrspur vorausfahrende Fahrzeug 2 sowie ein auf einer benachbarten Fahrspur befindliches Fahrzeug 3. Ziel des nachfolgend näher beschriebenen Objektdetektionsverfahrens ist es vor allem, selbsttätig die Lage dieser beiden Objekte 2, 3 bezogen auf das eigene Fahrzeug rasch und zuverlässig zu ermitteln.

Hierfür wird jedes der beiden Stereobilder 1a, 1b in ein entsprechendes Strukturklassenbild mittels einer Strukturklassifizierung transformiert, wie sie in Fig. 2 schematisch angedeutet ist. Für jeden Bildpunkt x der beiden Stereobilder 1a, 1b wird eine Umgebungsbildpunktmenge definiert, die im Fall von Fig. 2 aus den vier in Zeilen- und Spaltenrichtung, d. h. parallel und senkrecht zur horizontalen Epipolarlinie, nächstbenachbarten Umgebungsbildpunkten x_0, x_1, x_2, x_3 besteht. Wenn sich diese Wahl der Umgebungsbildpunkte als zu lokal erweist, kann alternativ ein größerer Nachbarschaftsbereich von Umgebungsbildpunkten bzw. ein Satz weiter entfernter Umgebungsbildpunkte gewählt werden. Bei Wahl eines größeren Nachbarschaftsbereiches, die eine kubische Steigerung der Anzahl von Strukturklassen impliziert und die Empfindlichkeit bezüglich der jeweils gewählten Helligkeitsschwellwerte anwachsen läßt, ist die Auswertung in einer Gaußpyramide von Vorteil. Zufriedenstellende Ergebnisse werden durch eine rechen-technisch weniger aufwendige Erhöhung der Abgriffweite erzielt, d. h. anstelle der in Fig. 2 gewählten, in Zeilen- und Spaltenrichtung direkt benachbarten Umgebungsbildpunkte werden dann solche gewählt, die

zum Bezugspunkt x einen Abstand von einem oder einigen wenigen Bildpunkten besitzen.

In jedem Fall wird dann paarweise die Helligkeit des Bezugspunktes x mit denjenigen des jeweiligen Umgebungspunktes x_0 bis x_3 unter Vorgabe eines positiven Helligkeitsschwellwertes T in digitalisierter Form gemäß einer Ternärlöge verglichen, wobei als Vergleichsergebnis ein Helligkeitsdifferenz-Digitalwert $d(x_i)$ in Form einer jeweiligen Ternärziffer gemäß nachfolgender Beziehung generiert wird:

$$d(x_i) = \begin{cases} 1, & \text{wenn } g(x_i) - g(x) < T \\ 2, & \text{wenn } g(x_i) - g(x) < -T, \\ 0, & \text{sonst} \end{cases}$$

wobei $g(x)$ die Helligkeit des Bezugspunktes und $g(x_i)$ mit $i=0, \dots, 3$ die Helligkeit des jeweiligen Umgebungspunktes ist. Die vier zu einem jeweiligen Bildpunkt x in dieser Weise ermittelten Ternärziffern $d(x_0)$ bis $d(x_3)$ werden dann zu einer Gruppe in Form einer Ternärzahl C der Form

$$C = \sum_{i=0}^3 3^i \cdot d(x_i)$$

zusammengefaßt, wie in der rechten Hälfte von Fig. 2 veranschaulicht. Für die so gebildete Ternärzahl C sind einundachtzig unterschiedliche Werte möglich, von denen jeder eine eigene Strukturklasse definiert. Anstelle der vorliegend beschriebenen ternären Digitalisierung der Helligkeitsunterschiede kommen auch andere Mehrwert-Digitalisierungen in Betracht, allerdings werden bei einer nur binären Digitalisierung Gradienteninformationen, die ein wichtiger Bestandteil der nachfolgenden Korrespondenzanalyse sind, unterdrückt, während höherwertige Digitalisierungen einen erhöhten Rechenaufwand zur Folge haben.

Fig. 3 zeigt das Ergebnis der Strukturklassifikation für das linke Stereobild 1a von Fig. 1, wobei als Gesamthelligkeitsumfang das Intervall $[0,255]$ und ein Helligkeitsdifferenzschwellwert von $T=20$ gewählt wurden. Zur besseren Veranschaulichung wurden die ermittelten Ternärzahlen C mit dem Faktor drei multipliziert und invertiert, so daß die zur Strukturklasse mit der Ternärzahl $C=0$ gehörigen Bildpunkte in Fig. 3 weiß erscheinen. Aus Fig. 3 läßt sich erkennen, daß aufgrund der durch den Helligkeitsschwellwert T eingebrachten Nichtlinearität eine beträchtliche Datenreduktion erzielt wird. Im Bild von Fig. 3 gehören etwa 50% aller Bildpunkte der Strukturklasse mit $C=0$ an. Weitere 25% der Bildpunkte gehören zu sogenannten horizontalen Strukturklassen, das sind jene mit horizontalen Strukturen, bei denen $d(x_3)=d(x_1)=0$ ist. Solche horizontalen Strukturklassen sind für die Auswertung von Stereobildern mit der vorliegend horizontal gewählten Epipolarlinie nicht nutzbringend und werden daher bei der weiteren Bildauswertung nicht mehr berücksichtigt. Demgemäß brauchen nur noch 25% der Bildpunkte einer Korrespondenzanalyse unterworfen werden. Durch Auswahl eines geeigneten Auswertefensters läßt sich diese Bildpunktanzahl unter Erhöhung der Auswertesicherheit weiter steigern. So kann z. B. im vorliegenden Fall der merklich über dem Horizont und damit der

Straßenoberfläche liegende Bildbereich als nicht interessierend ausgeblendet werden. Ein günstiges Auswertefenster ist beispielhaft im linken Stereobild 1a von Fig. 1 mit dem dortigen weißen Rahmen 4 angedeutet. Mit der Wahl dieses Auswertefensters 4 wird bei der vorliegenden Bildszene die Menge an für die anschließende Korrespondenzanalyse zu berücksichtigenden Bildpunkten auf 13% aller Bildpunkte reduziert, und zugleich werden Störeinflüsse durch Disparitätsfehlbestimmungen vermieden, die auf ausgeblendete Bildbereiche zurückgehen.

An die Erzeugung des Strukturklassenbildpaars schließt sich eine vereinfachte Korrespondenzanalyse an, die zwar keine bildpunktweise eindeutige Lösung des Korrespondenzproblems darstellt, jedoch mit entsprechend geringerem Rechenaufwand ausreichend zuverlässige Ergebnisse für die hier in Betracht kommenden Anwendungen liefert, bei denen die relevanten, zu detektierenden Objekte Strukturen mit zur Epipolarlinie im wesentlichen senkrechten Helligkeitsgradienten besitzen, was nicht nur für Fahrzeuge, sondern auch für die meisten stehenden Objekte und insbesondere für Fußgänger gilt. Dazu wird im einzelnen wie folgt vorgegangen.

Zu jedem Bildpunkt im einen Strukturklassenbild als Bezugspunkt wird innerhalb eines vorgegebenen Disparitätsintervall nach auf der entsprechenden Epipolarlinie als der maßgeblichen Suchrichtung liegenden Bildpunkten derselben Strukturklasse im anderen Strukturklassenbild gesucht. Wegen der parallelen optischen Achsen der Bildaufnahmekameras und der horizontal versetzten Anordnung derselben ergibt sich im vorliegenden Beispiel eine in Bildzeilenrichtung verlaufende Suchrichtung. Die Entfernung eines jeweils aufgefundenen Bildpunktes zur Lage des Bezugspunktes wird erfaßt und als zugehöriger Disparitätswert in einem Disparitätshistogramm akkumuliert, und zwar in einer ersten Alternative ungewichtet, d. h. die Häufigkeit im Histogramm wird für jeden Disparitätswert pro Auftreten desselben um eins erhöht. In einer zweiten Alternative wird ein den Bildkontrast berücksichtigendes, gewichtetes Häufigkeitsinkrement verwendet. Dazu wird für jeden Bildpunkt x ein lokales Kontrastmaß aus den ermittelten Helligkeitsdifferenzen generiert, z. B. ein Kontrastmaß $K(x)$ der Form

$$K(x) = \frac{1}{N} \sum_i |g(x_i) - g(x)|,$$

wobei N die Anzahl der Umgebungspunkte angibt, deren Helligkeit um mehr als der Helligkeitsschwellwert T von der Bezugspunkthelligkeit abweicht, und bei der Summation auch nur diese Umgebungspunkte berücksichtigt werden. Damit erhält der einzelne Bildpunkt ein Kontrastgewicht, das es erlaubt, kontrastreiche Strukturen bei der Disparitätsanalyse zu betonen. Eine Möglichkeit, dieses Kontrastmaß in der Erstellung des Disparitätshistogramms zu berücksichtigen, besteht darin, den zu einem jeweiligen korrespondierenden Bildpunkt ermittelten Disparitätswert nicht mit dem Häufigkeitsinkrement eins, sondern mit einem um das Minimum der zu den beiden korrelierenden Bildpunkten gehörigen Kontrastwerte erhöhten Häufigkeitsinkrement in die Histogrammerzeugung eingehen zu lassen. Dies ergibt eine schärfere Trennung der Häufigkeitspunktbereiche, d. h. der Peakbereiche, im resultierenden Disparitätshistogramm vom Untergrundrauschen.

schen:

Fig. 4 zeigt das durch den erläuterten Disparitätsanalysevorgang unter Verwendung des ungewichteten Häufigkeitsinkrementes eins aus dem Bereich des Auswertefensters 4 erhaltene Disparitätshistogramm mit Disparitätswerten zwischen null und zweihundert. Dabei wurden zu jedem Bezugspunkt im einen Strukturbild alle in Richtung der entsprechenden Epipolarlinie liegenden Bildpunkte im anderen Strukturbild berücksichtigt. Alternativ dazu hat sich für den vorliegenden Anwendungsfall der Beobachtung des Straßenverkehrs die Favourisierung kleiner Disparitätswerte als vorteilhaft erwiesen, was mit einer Bevorzugung großer Entfernungen korreliert. Denn in komplexen Straßenszenen bedingen häufig periodische Strukturen, wie z. B. die Häuserfront 5 im Stereobildpaar 1a, 1b von Fig. 1, ein starkes Rauschen im Disparitätshistogramm, indem sie wegen ihrer Periodizität sehr nahe am Kameraort stehende Phantomobjekte vortäuschen. Dem wird durch die Favourisierung großer Entfernungen wirksam begegnet. In einer einfachen und effektiven Form geschieht dies dadurch, daß bei der Disparitätsanalyse zu jedem Bezugspunkt im einen Strukturklassenbild nur der diesem auf der Epipolarlinie nächstgelegene Bildpunkt derselben Strukturklasse im anderen Strukturklassenbild berücksichtigt wird. Diese Vorgehensweise erweist sich für die Beobachtung des Straßenverkehrs bei den praktisch auftretenden Straßenszenen mit ihren relativ großen Hindernissen als zulässig und wegen der dadurch bedingten Einsparung an Rechenzeit und der Unterdrückung von Phantomobjekten sehr vorteilhaft. Alternativ zu dem in Fig. 4 gezeigten zweidimensionalen Histogramm kann bei Bedarf ein dreidimensionales Disparitätshistogramm als quantisierte Tiefenkarte angefertigt werden, bei dem als zusätzliche Histogrammachse der laterale Offset des jeweiligen Bezugspunktes zur Bildmittellachse herangezogen wird.

Nachdem das Disparitätshistogramm in der einen oder anderen Form erstellt wurde, wird dieses auf Häufungspunktbereiche hin untersucht. In Fig. 4 sind von den ausgeprägten Häufigkeitspunktbereichen die zwei d_1 , d_2 mit den höchsten Disparitätswerten markiert, wobei der eine, d_1 , einem Disparitäts-Spitzenwert von etwa 28 und der andere, d_2 , einem Disparitäts-Spitzenwert von etwa 15 entspricht. Jeder solche Häufigkeitspunktbereich sollte einem Objekt im Beobachtungsbereich entsprechen. Um dies zu verifizieren und dadurch die Objektdetektion zu vervollständigen, werden anschließend in zur Disparitätsanalyse umgekehrter Richtung diejenigen Bildpunkte eines Strukturklassenbildes wiedergegeben, die zu einem jeweiligen Disparitätshäufungspunktbereich gehören. So sind in Fig. 5 die zu dem Disparitätshäufungspunktbereich d_1 mit dem Disparitäts-Spitzenwert 28 gehörigen Bildpunkte und in Fig. 6 die zu dem Disparitätshäufungspunktbereich d_2 mit dem Disparitäts-Spitzenwert 15 gehörigen Bildpunkte dargestellt. Wie daraus zu erkennen, gibt Fig. 5 bereits recht gut das auf der gleichen Fahrspur vorausfahrende Fahrzeug 2 wieder, während das auf der anderen Fahrspur befindliche Fahrzeug 3 das Bild von Fig. 6 dominiert. Es kann daraus geschlossen werden, daß diese beiden Fahrzeuge 2, 3 zu diesen beiden Disparitätshäufungspunktbereichen d_1 , d_2 gehören.

Aus dieser Zuordnung eines endgültigen Disparitätswertes zu einem jeweiligen Objekt kann die Entfernung L des letzteren von der Stereokameraanordnung durch die Beziehung $L = f_x \cdot B/d$ ermittelt werden, wobei B den

Basisabstand der Stereokameraanordnung, f_x deren skalierte Brennweite und d den jeweiligen Disparitätswert bezeichnen. Anschließend kann das detektierte Objekt bei Bedarf kontinuierlich weiter beobachtet, d. h. verfolgt werden. Dazu wird einleitend eine einfache Ballungsanalyse vorgenommen, bei der ein rechteckförmiges Fenster minimaler Höhe und Breite so im resultierenden Objektbild analog den Fig. 5 und 6 positioniert wird, daß darin ein vorgegebener Mindestprozentsatz von z. B. 90% der gefundenen Bildpunkte enthalten ist. Da die obige trigonometrische Beziehung für die Objektentfernung L wegen der Diskretisierung der Disparitätswerte vor allem bei kleinen Disparitäten eine relativ grobe Näherung darstellt und in diesem Fall für die Bestimmung von Bewegungsvorgängen noch nicht optimal ist, kann es zudem zweckmäßig sein, eine verbesserte, einer subbildpunktgenauen Messung entsprechende Disparitätsschätzung vorzunehmen, bei der zum Spitzenwert eines jeweiligen Disparitätshäufungspunktbereiches ein Korrekturterm der Form

$$d_k = 0,5 \cdot (h_+ + h_-) / (2h_m - (h_+ + h_-))$$

hinzugaddiert wird, wobei h_m den Spitzenwert und h_+ bzw. h_- die Disparitätshöhen im jeweiligen Seitenbereich bezeichnen.

Nach der solchermaßen bewerkstelligten Objektdetektion mit Bestimmung der Entfernung und der ungefähren Ausdehnung des jeweiligen Objektes wird in der nachfolgenden Verfolgungsphase eine auf den relevanten, gerade noch das zu beobachtende Objekt enthaltenden Bildbereich beschränkte Bildauswertung nach dem beschriebenen Verfahren in zyklisch wiederholter Form durchgeführt. Die erfindungsgemäßen Verfahrensmaßnahmen ermöglichen typische Verfahrenszyklen von lediglich größenordnungsmäßig zwischen 100ms und 200ms, verglichen mit typischen Zykluszeiten aufwendigerer Stereobildauswerteverfahren von mehreren Sekunden. Durch die Bildbereichsbeschränkung wird nicht nur Rechenzeit eingespart, sondern vor allem auch der Einfluß störender Hintergrundobjekte reduziert. Von zentraler Bedeutung ist, daß nur noch ein relativ kleiner Disparitätsbereich ausgewertet werden muß, indem um das zu verfolgende Objekt gedanklich ein Kubus gelegt wird, dessen Tiefe von der maximal zu erwartenden Entfernungsänderung abhängt.

Die laterale Position eines interessierenden Objektes läßt sich für die Spurverfolgung mittels Strahlensatz beispielsweise aus der lateralen Position des Bildmittelpunktes oder alternativ des Bildschwerpunktes ermitteln. Zweckmäßigerweise werden die erhaltenen Resultate einer üblichen Kalman-Filterung unterworfen, um durch Berücksichtigung des Wissens über die eingeschränkten Bewegungsmöglichkeiten zu optimalen Resultaten zu gelangen.

Um Straßenoberflächenstrukturen von den Strukturen von über der Straßenoberfläche befindlichen Objekten zu isolieren, kann ein einfaches Straßenmodell Verwendung finden, bei dem eine ebene Straßenoberfläche und parallele optische Achsen der Stereokameraanordnung angenommen werden. Für die zu erwartende Disparität $d_g(x,y)$ eines Bildpunktes mit den Koordinaten (x,y) in der Bildebene ergibt sich dann aus der Kamerahöhe H über der Fahrbahn und dem Kameranigungswinkel α gegenüber dem Horizont die Beziehung

$$d_g(x,y) = (B/H) f_x ((y/f_y) \cos \alpha + \sin \alpha),$$

wobei f_x und f_y die skalierten Brennweiten in den betreffenden Koordinatenrichtungen und B die Basisweite der Stereokameraanordnung bezeichnen. Zur Extraktion der Straßenoberflächenstrukturen tragen unter diesen Voraussetzungen nur Punkte unterhalb des Horizontes bei. Dabei liegen alle Bildpunkte mit Disparitätswerten, die innerhalb eines bestimmten Intervalls um den Erwartungswert liegen, auf der Straßenoberfläche, während diejenigen mit größeren Disparitätswerten zu Objekten über der Straßenoberfläche gehören und diejenigen mit zu kleinen Disparitätswerten von Reflexionen weiter entfernter Objekte herrühren.

Je nach Anwendungsfall können für die Objektdetektion zusätzlich Symmetrieüberlegungen verwendet werden, die den Rechenaufwand verringern und die Detektionssicherheit erhöhen können. In jedem Fall ermöglicht das erfindungsgemäße Verfahren eine bei gegebener Detektionssicherheit sehr rasche und robuste Objektdetektion und läßt sich zur optischen Echtzeit-Objektbeobachtung im Straßenverkehr, auf dem Gebiet der Robotik und überall dort einsetzen, wo Bedarf an einer stereobildbasierten Detektion und gegebenenfalls Verfolgung von Objekten besteht.

Patentansprüche

1. Verfahren zur Stereobild-Objektdetektion, insbesondere für Straßenfahrzeuge, gekennzeichnet durch folgende Schritte:

- a) Erzeugen eines Strukturklassenbildpaares aus dem jeweiligen aufgenommenen Stereobildpaar (1a, 1b), indem für jeden Bildpunkt (x) die Differenzen zwischen dessen Helligkeitswert ($g(x)$) und jedem der Helligkeitswerte ($g(x)_i, i=0, \dots, n-1$) einer Anzahl (n) vorgegebener Umgebungsbildpunkte (x_0, \dots, x_{n-1}) als Digitalwerte ($d(x_i)$) bestimmt und die erhaltenen Digitalwerte zu einer Digitalwertgruppe (C) zusammengefaßt werden, wobei jeweils identische Digitalwertgruppen eine eigene Strukturklasse definieren,
- b) Durchführen einer Korrespondenzanalyse des Strukturklassenbildpaares, bei der
 - b. 1) nur diejenigen Strukturklassen berücksichtigt werden, deren Bildpunkte wenigstens einen in Richtung der Epipolarlinie liegenden Umgebungsbildpunkt mit um eine oder mehrere Helligkeitsdigitalstufen unterschiedlicher Helligkeit besitzen,
 - b. 2) für jeden Bildpunkt einer jeweils noch berücksichtigten Strukturklasse des einen Strukturklassenbildes die innerhalb eines vorgebbaren Disparitätsintervalls auf der Epipolarlinie liegenden Bildpunkte derselben Strukturklasse des anderen Strukturklassenbildes aufgesucht und der zugehörige Disparitätswert ermittelt wird und
 - b. 3) die dadurch insgesamt erhaltenen Disparitätswerte entsprechend einem zugeordneten Häufigkeitsinkrement in einem Disparitätshistogramm akkumuliert werden, und
 - c) Identifizieren von Häufungspunktbereichen (d_1, d_2) im erzeugten Disparitätshistogramm und Extrahieren eines jeweiligen Objektes aus der zum jeweiligen Häufungspunktbereich gehörenden Bildpunktgruppe.

2. Verfahren nach Anspruch 1, weiter dadurch gekennzeichnet, daß zur Digitalwertermittlung im

Schritt a ein Ternärsystem verwendet wird, wobei den Helligkeitsdifferenzen, die betragsmäßig kleiner gleich einem vorgegebenen positiven Helligkeitsschwellwert (T) sind, eine erste Ternärziffer, den Helligkeitsdifferenzen, die größer als der Schwellwert sind, eine zweite Ternärziffer und den Helligkeitsdifferenzen, die kleiner als das Negative (-T) des Schwellwertes sind, die dritte der drei Ternärziffern zugeordnet wird.

3. Verfahren nach Anspruch 1 oder 2, weiter dadurch gekennzeichnet, daß als Umgebungsbildpunkte (x_0 bis x_3) jeweils zwei auf gegenüberliegenden Seiten des jeweiligen Bezugsbildpunktes (x) zum einen parallel und zum anderen senkrecht zur Epipolarlinie liegende Bildpunkte verwendet werden, die dem Bezugsbildpunkt direkt benachbart oder von diesem um eine vorgebbare Abgriffweite beabstandet angeordnet sind.

4. Verfahren nach einem der Ansprüche 1 bis 3, weiter dadurch gekennzeichnet, daß das vorgebbare Disparitätsintervall für die Korrespondenzanalyse so gewählt wird, daß zum jeweiligen Bildpunkt im einen Strukturklassenbild nur der nächstliegende Bildpunkt auf der Epipolarlinie im anderen Strukturklassenbild berücksichtigt und der diesbezügliche Disparitätswert im Disparitätshistogramm akkumuliert wird.

5. Verfahren nach einem der Ansprüche 1 bis 4, weiter dadurch gekennzeichnet, daß jedem Bildpunkt des Strukturklassenbildpaares zusätzlich zur jeweiligen Digitalwertgruppe ein in Abhängigkeit von den Helligkeitsdifferenzen ermittelter Kontrastwert ($K(x)$) als Gewichtungsfaktor zugeordnet wird und der zu einem jeweiligen Bildpunkt ermittelte Disparitätswert mit einem in Abhängigkeit von den zugehörigen Kontrastwerten gewichteten Häufigkeitsinkrement zum Disparitätshistogramm akkumuliert wird.

6. Verfahren nach einem der Ansprüche 1 bis 5, weiter dadurch gekennzeichnet, daß die Extraktion des jeweiligen Objektes aus der zu einem jeweiligen Disparitätshäufungspunktbereich gehörenden Bildpunktgruppe eine Ballungsanalyse der Bildpunktgruppe umfaßt, mit der ein Rechteck minimaler Fläche ermittelt wird, das einen vorgebbaren Prozentsatz der Bildpunkte dieser Bildpunktgruppe enthält.

7. Verfahren nach einem der Ansprüche 1 bis 6, weiter dadurch gekennzeichnet, daß nach erstmaliger Extraktion des jeweiligen Objektes aus der zu einem jeweiligen Disparitätshäufungspunktbereich gehörenden Bildpunktgruppe eine Bewegungsverfolgung desselben mittels wiederholtem Ausführen der Schritte a bis c eingeschränkt auf den Bereich eines minimalen, das betreffende Objekt enthaltenen Kubus durchgeführt wird.

8. Verfahren nach einem der Ansprüche 1 bis 7 zur Detektion von Objekten im Straßenverkehr, weiter dadurch gekennzeichnet, daß unter der Annahme eines im wesentlichen ebenen Straßenverlaufs und paralleler optischer Stereokameraachsen diejenigen Bildpunkte, deren zugeordnete Disparitätswerte innerhalb eines vorgegebenen Erwartungsbereichs um einen der folgenden Beziehung genügenden Disparitätserwartungswert (d_g)

$$d_g(x, y) = (B/H) f_x ((y/f_y) \cos \alpha + \sin \alpha),$$

Hierzu 2 Seite(n) Zeichnungen

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55
 60
 65

1. *Journal of the American Medical Association*, 1990; 263: 1025-1028.

- Leerseite -

Figure 1 consists of two scatter plots. The left plot shows a positive correlation between the number of children and the number of adults, with a regression line indicating a positive slope. The right plot shows a negative correlation between the number of children and the number of adults, with a regression line indicating a negative slope.

100

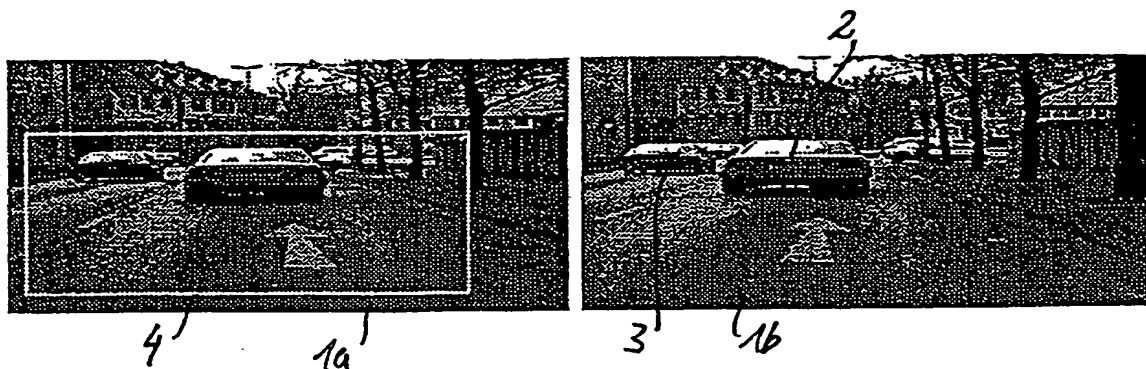


Fig. 1

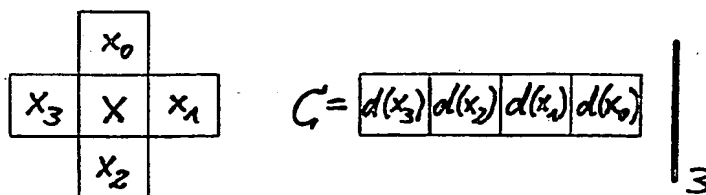


Fig. 2

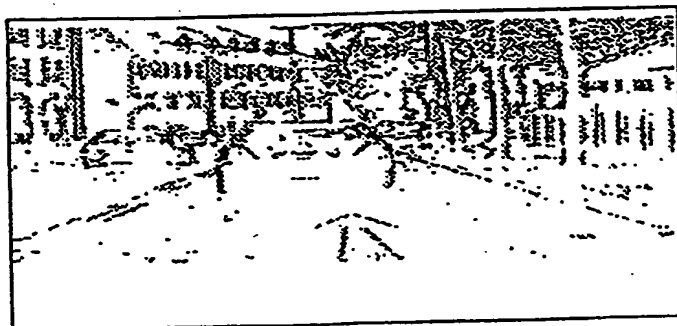


Fig. 3

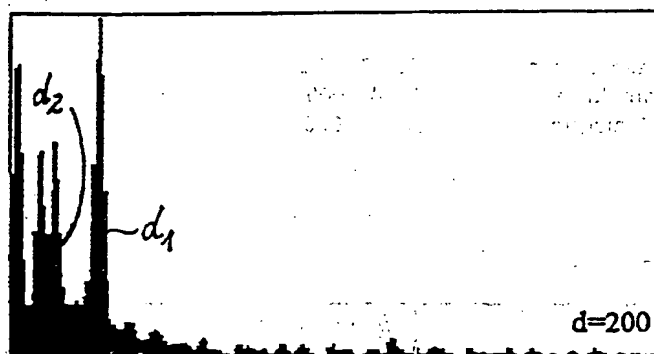


Fig. 4

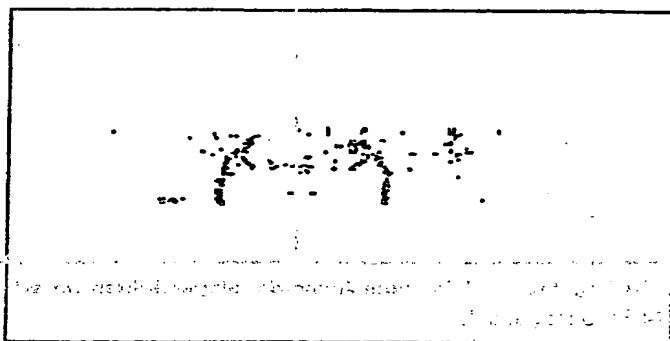


Fig. 5

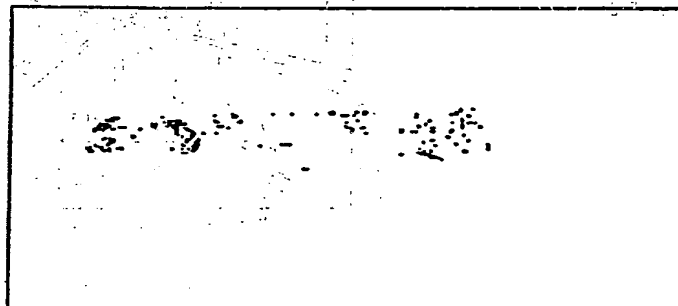


Fig. 6